# Price Discount and Price Dispersion in Online Market: Do More Firms Still Lead to More Competition? 

## Wenche Wang ${ }^{1}$, Fan Li ${ }^{2}$ and Yujia Zhang ${ }^{3}$

${ }^{1}$ University of Michigan, School of Kinesiology, Ann Arbor, USA, wwenche@umich.edu
${ }^{2}$ Shenzhen University, China Center for Special Economic Zone Research, Shenzhen, China, lifan@szu.edu.cn
${ }^{3}$ Shenzhen University, China Center for Special Economic Zone Research, Shenzhen, China, zhangyujia2018@email.szu.edu.cn

Received 2 December 2019; received in revised form 14 July 2020; accepted 24 July 2020


#### Abstract

Contrary to the traditional economic prediction that prices decrease as more firms enter the market, this paper finds a non-monotonic relationship between the number of firms and prices in online book markets. A substantial decrease in posted prices is observed when the market moves from a monopoly to duopoly. Surprisingly, as the market further expands, posted prices converge to monopoly price rather than competitive price. This is because as the number of sellers increase on the internet, the cost for consumers to obtain accurate information as well as the cost to evaluate products and sellers increase. Thus, the existence of a large number of sellers deter consumer search and lead to more random purchases. With little consumer search, sellers' returns to undercutting price reduce. Therefore, instead of lowering price when faced with more competitors, sellers increase prices.


Keywords: Competition, Market structure, Online market, Pricing strategy, Search

## 1 Introduction

The rapid growth in the use of the Internet over the last few decades has been driving consumers and sellers to encompass more online transactions. Compared to searching across brick-and-mortar stores, it is much easier for consumers to obtain price information in the online market. As the number of sellers increases online, if consumers have perfect and complete information, price competition is expected to increase, leading to perfect competition and price convergence. In other words, the increase in the number of sellers would reduce both price and price dispersion.

There have been studies that explore the effect of information asymmetry in the internet market and challenge the argument of frictionless search. For example, Ellison and Ellison [21] suggest that although the internet facilitates price search, it also allows firms to adopt a number of strategies to complicate search. In particular, they found that online retailers may show lower priced products on price comparison websites to attract consumers but show higher quality and thus higher margin products once consumers arrive on their websites. Hossain and Morgan [26] and Blake et al. [8] find that online sellers often don't display shipping fees until consumers reach the final purchase page. Generally, unlike shopping at traditional brick-and-mortar stores, online shopping faces greater product and service uncertainties as consumers must determine their preference without observing product quality. The costs for consumers to evaluate products and compare stores increase as more firms enter market. Meanwhile, the costs for sellers to observe market demand and rivals' prices are significantly reduced. Moreover, price adjustment becomes almost costless on the internet. Thus, sellers can conduct more frequent price adjustments and adopt various promotional strategies that increase consumer search cost [14], [28].

Motivated by these findings in the literature, the current study explores online sellers' strategic behavior and the relationship between the number of firms and market competitiveness. A unique dataset of store level and product level information as well as prices and aggregate sales for new fiction books is obtained from Tmall (Site 1), the largest Business-to-Consumer (B2C) online platform in China. We seek to answer two questions. First, we investigate whether it is indeed less costly for consumers to gather information to make purchasing decisions on the internet which include produce and service information as well as prices. Second, we would like to examine the relationship between the number of firms and market competitiveness under such information structure in the internet market.

We find that in online markets, though the cost of obtaining price information is significantly reduced, the cost for consumers to gather product information as well as store information has not changed. This cost may even increase when stores provide partial information or even contradicting information about the products. Contrary to standard economic predictions, we find that prices do not always decrease as the number of firms carrying the products increases. More firms can introduce competition, but can also make consumer search more costly. The effect of competition and the effect of consumer search on prices are offsetting. Competition effect only dominates when a market changes from monopoly to duopoly. When more firms enter the market, searching through every store becomes too costly. When consumers cannot search through all the stores, their purchase decisions become less predictable as they depend on the random set of stores that consumers search through. The expected benefit for firms to undercut price reduces. Therefore, instead of lowering prices, firms will raise prices. In the meantime, we find that price dispersion increases as the number of firms increase which is consistent with the predictions from the search theory and the empirical evidence found in the literature [13], [25], [27].

The organization of the rest of the paper is as follows. Section 2 reviews related literature. Section 3 provides an introduction of the data as well as detailed descriptive statistics. Section 4 presents our hypotheses and empirical models. Section 5 discusses our results and section 6 concludes the paper.

## 2 Literature Review

Traditional models of competition with perfect information cannot explain the widely observed price dispersion in the real world as information affects the effective degree of competition [43]. Information asymmetry has long been recognized to make markets behave less competitively [42], [40]. When consumers have to incur a positive cost to obtain information of the product, search theory raises serious concerns over the conventional result that increasing the number of firms would increase the competitiveness of the market [19], [44]. On the internet market, however, price and product information can be obtained with a simple search, thus creating a seemingly perfect information environment. In the economic and business literature, there is a debate over whether there exists frictionless search on the internet [14], [21], [28]. When consumers' cost of information acquisition is very low or even zero, the effective degree of competition would be determined by the number of firms. The impact of search cost and subsequently competition will be reflected in realized prices and price dispersion [39]. Thus, a large body of the literature focus on the comparison between online and offline prices and the theoretical and practical explanations to their empirical evidence on online prices and online price dispersion.

### 2.1 Online Price vs. Offline Price

The development of Internet technologies proliferates the growth of internet markets. Internet markets differ from brick-and-mortar markets in their information structure which affects consumer search behavior and subsequently price and price dispersion. It is widely believed that compared to physical markets, it is easier to gather and compare information for potential transactions in online markets. The internet also provides lower entry cost and transaction costs which would lead to a convergence to the economic ideal of perfect competition [1], [5]. Such reductions in costs are expected to drive down both price and price dispersion.

A number of empirical studies have found lower online prices from various markets. Brynjolfsson and Smith [11] was the first to examine online prices and found substantially lower prices of books and CDs on the internet. This reduction in prices on the internet is further confirmed by Brown and Goolsbee [10] in the insurance markets, Pan et al. [35] in CDs, DVDs, desktop and laptop computers, Morton et al. [33] in automotive sales, and Orlov [34] in the airline industry. However, there are also some studies that suggest the contrary. Bailey [4] finds higher prices on the internet using a sample of books, CDs, and software. Pan et al. [35] find similar results in their study of books and software. Ancarani and Shankar [2] suggest online prices are indeed higher than brick-and-mortar prices when accounting for shipping costs in the Italian book and CD market. Erevelles et al. [22] study the vitamin prices and again find significantly higher prices on the internet. Meanwhile, there are also studies that document no significant difference between prices online and offline. For example, Clay et al [17] find that online book prices do not differ from offline book prices significantly. Pan et al. [35] also find similar prices of electronic products and personal digital assistant products online and offline.

While the empirical evidence on the direct comparison between online and offline prices appear to be mixed, the findings on online price dispersion generally suggest a consistent story - substantial price dispersion remains on the internet. Empirical evidence of a large and persistent online price dispersion has been found in the book market [11], airline markets [18], [34], and electronic products [6]. Pathak [37] found that even with online comparison shopping agents, there still exists price dispersion due to selection issue by the agents and temporal delay in updating information.

### 2.2 Explanation to Empirical findings on Online Price and Price Dispersion

Given the mixed results in online prices and the evidence of persistent online price dispersion, researchers turn to explore the reasons behind the complicated empirical evidence on online prices and to explain why price dispersion does not disappear when online price search is almost costless.

Two main explanations have been provided. First, online retailers may provide different qualities of services that are not observed by researchers. Second, the frequent promotions that involve price adjustments by online sellers may increase the cost of a price search. Additionally, the large amount and various product and service information provided by the retailers also affect the cost for consumers to identify their ideal product and retailer.

While consumers can obtain information on product characteristics with very low cost on the internet, it is not always easy to compare the service between retailers. Service quality such as shipping policy, return policy, delivery speed, responses to consumer questions and requests may differ substantially across retailers [30]. Retailers that offer higher quality service can develop strong brands and charge higher prices [47]. Thus, the unobserved heterogeneity in seller characteristics such as service quality, consumer awareness of the sellers, and reputation, may explain the mixed results between online and offline prices as well as the persistent online price dispersion [3], [37].

Another stream of explanations relies on sellers' strategic behavior that may affect consumer search costs. Unlike in the physical market, online retailers can adjust prices with much greater flexibility and with significantly lower costs [9], [12], [31]. Online retailers' frequent change in prices make it costly for consumers to conduct price comparison.

When consumers search for a product online, they usually are not conducting a simple price comparison of the same product across sellers. They typically have to research on product characteristics, service quality, retailer reputation, and lastly prices of the product as well as any additional charges for the service. This multi-dimensional evaluation of information provides retailers opportunity to shape consumer search process [46]. For example, Lynch and Ariely [32] found that if price was available on the first page, consumers would focus on price; if consumers needed to click further to obtain price information, other attributes would play a bigger role in affecting consumers' purchasing decisions. Rather than zero search cost online, there is a positive cost for consumers to make an extra click during their search [23], [24], [41]. As a result, consumers may stop searching sooner than what is expected under the assumption of zero search cost. This positive cost opens up an opportunity for online retailers to influence the search process and the information received and processed by consumers. Hossain and Morgan [26] and Blake et al. [8] both show that retailers would not reveal shipping cost until the last stage of the purchase. Dinerstein et al. [20] suggest that the design of the search algorithm on eBay affects the markups charged by the retailers.

In this paper, we investigate online sellers' strategic behavior using a dataset of new fiction books on Tmall (Site 1). Building upon the existing findings that sellers have an ability to obfuscate and make consumer search more costly on
the internet, we examine sellers' strategic information release which may complicate consumers' search process. If consumer search cost on the internet is endogenous on seller behaviors, we then ask if more sellers in the market would increase market competition, which is traditionally assumed in physical markets. We analyze the impact of the number of sellers on both price and price dispersion and find striking evidence that market competitiveness on the internet is not monotonic in the number of competitors. While our findings provide more evidence to support the existing literature's challenge on frictionless search on the internet, we also cast a new question to the study of internet market about competition in the presence of information asymmetry and sellers' strategic behavior.

## 3 Data

We collect posted information on new fiction books between November 14 ${ }^{\text {th }}$ and December $15^{\text {th }} 2013$ from Tmall (Site 1), commonly referred to as Tmall, the largest business-to-consumer (B2C) online platform operated by Alibaba Group. Book has become one of the largest online retail categories [15] and the book market in China has gradually shifted from brick-and-mortar market to online market. Thus, we choose books as our subject of study. In addition, books are standardized items that can be identified by ISBN. We restrict our attention to new books for quality control purposes; additionally, we exclude textbooks from the dataset because consumers have a different demand structure for textbooks than for other books. Thus, we can identify the impact of information asymmetry and consumer search on the prices and price dispersions, avoiding the impact of product differentiation.

### 3.1 Descriptive Statistics

Our dataset contains 77,071 observations of book-store pairs that cover 21,065 book titles and 901 online bookstores. In our dataset, some bookstores reported an invalid ISBN and a few others did not report publisher's recommended prices. Some listings are bundles of several books and, therefore, hard to assess prices of individual books in the bundles. These observations were excluded from our sample and our sample is reduced to 68,382.

### 3.1.1 Sellers' Information Revelation

Similar to eBay, Tmall allows sellers to self report and organize book information. While this allows sellers to customize its own retail page, it also creates the potential for sellers to report different information for the same book. Thus, consumers may be exposed to different or even contradicting information on the same book. Generally, sellers can provide information on catalog, editor's recommendation, author introduction, and content summary for the book. On average, each listing only provides two out of the four categories of book information. Only about 10\% of the listings report all four information.

Table 1: Summary statistics and correlation of sellers' self-reported book information

|  | Mean | Catalog | Editor's <br> Recommendation | Author <br> Introduction | Content <br> Summary |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Catalog | 0.71 | 1.00 |  |  |  |
|  | $(0.25)$ |  |  |  |  |
| Editor's Review | 0.39 | 0.38 | 1.00 |  |  |
|  | $(0.30)$ |  |  |  |  |
| Author Introduction | 0.65 | 0.62 | 0.44 | 1.00 |  |
|  | $(0.27)$ |  |  |  |  |
| Content Summary | 0.30 | 0.09 | 0.16 | 0.02 | 1.00 |
|  | $(0.31)$ |  |  |  |  |
| Number of Observation | 9,911 |  |  |  |  |

Note: 1. Data is extracted from Tmall (Site 1).
2. Observations are measured in book level and each variable represents the proportion of listings that provide each category of book information.
3. Books that are only available at one store are dropped from the sample.
4. Standard deviation of the average percentage of reported information is in parentheses.

Table 1 explores the self-reported information in book level. Each observation identifies a book and each variable represents the proportion of listings that provide each category of book information. While most stores provide information on catalog and author, less than half of the stores offer editor's recommendation or content summary. Note that this is a conservative measure of the heterogeneity of the information provided by each store since stores may even post different information under each category for a book. For Editor's Review, in particular, stores often post reviews from different editors. The correlation matrix suggests a very small percentage of stores providing content summary when they post any of the other three information. This indicates that sellers select product information to release. Whether such information revelation is selected intentionally or unintentionally, it creates significant
information asymmetry in the online market. The variation in the availability of information substantially increases the cost for consumers to compare the same product across stores.

### 3.1.2 Price and Discount

Table 2 summarizes the average prices and discounts of the books in our dataset. The average publisher's recommended price for the entire sample is 34.35 Chinese Yuan, which is equivalent to about 5.6 US dollars. Sales price is the listing price at the time of data generation. A posted price discount is defined as the difference between the publisher's recommended price and the sales price divided by the publisher's recommended price. The average discount in the sample is close to $30 \%$. Each listing reports the total sales of the book from the store as well as sales in the past month. We aggregate sales of each book across the stores and further separate books into two categories, books with positive sales and books with zero sales in the past month. The positive sales sample includes all books with sales in at least one store. These books are on average offered at a slightly higher discount than books without sales. In the full sample, each book is available at about 3.67 bookstores while the books with positive sales in the past month are carried by about 4.36 stores.

Table 2: Summary statistics of posted prices and discounts

|  | Full Sample | Books with <br> Positive Sales | Books with <br> Zero Sales |
| :--- | :--- | :--- | :--- |
| Average Sales Price | 23.97 | 24.67 | 23.30 |
| Average Publisher's Recommended Prices | 34.35 | 39.90 | 31.15 |
| Average Discount (\%) | 29.52 | 30.52 | 28.59 |
| Average Number of Stores per Book | 3.67 | 4.36 | 3.19 |
| Number of Book Title | 15,145 | 5,542 | 9,603 |
| Number of Bookstores | 752 | 740 | 346 |
| Number of Observations | 68,382 | 33,172 | 35,210 |

Note: 1. Data is extracted from Tmall (Site 1).
2. Price is measured in Chinese Yuan (RMB). According to Exchange-Rates.org, one US dollar was approximately 6.1 Chinese Yuan in the sample period.
3. Sales prices and discounts are summarized over all the listings in the sample while publisher's recommended prices are summarized over book titles.

In addition to book-level price discount, stores also conduct various store-level promotions, which could make price comparison for an individual book more difficult. Five types of promotions are observed in the data, conditioning on different amount of purchase. Table 3 summarizes such store-level discounts. A total of 103 different store level promotions conditioning on different amount of purchases are observed in the data. Some stores combine two or more promotions. On one hand, such store-level promotions make price comparison more costly for consumers; on the other hand, this leave sellers more flexibility to adjust price.

Table 3: Summary statistics for store level promotions

|  | Number of Conditions | Max Number of Promotion <br> (by Promotion Type) | Total Number of Stores <br> with Each Promotion |
| :--- | :--- | :--- | :--- |
| Free shipping | 25 | 2 | 70 |
| Gift with Purchase | 10 | 1 | 12 |
| Voucher | 19 | 5 | 25 |
| Price Deduction | 44 | 4 | 45 |
| Credit Reward | 5 | 2 | 5 |
| Number of Observation | 862 |  |  |

Note: 1. Data is extracted from Tmall (Site 1).
2. Observation is measured in store level.
3. Free shipping refers to store-wide free shipping. It does not include listings with $O$ for shipping cost.
4.Tmall assigns consumers credit scores which are based on sellers' rating on consumers after a transaction is completed.

Credit reward is a promotional strategy when sellers reward consumers additional credit score after a purchase.

### 3.1.3 Price Dispersion

Sellers conduct various forms of obfuscation and promotion strategies that can deter consumer search. Without much search from the consumers, posted price dispersion is expected to be large. Percentage price difference and coefficient of variation are used to measure the size of price dispersions. Percentage price difference is defined as $P D=$ $\frac{P^{\max }-P^{\min }}{\bar{P}}$, where $P^{\max }$ and $P^{\text {min }}$ are the maximum and minimum prices, respectively, and $\bar{P}$ is the average price. The coefficient of variation is defined as $C V=\frac{\sigma}{\bar{P}}$, where $\sigma$ is the standard deviation of prices. Percentage price difference measures the range of prices while coefficient of variation measures the volatility of prices. Percentage price difference puts entire weight on extreme prices. Some sellers may set extremely low prices on certain books in order to attract consumers into the store while increase prices of other books to their maximum levels, usually the publisher's recommended prices, in order to maximize profits. Thus, we further construct a 90th to 10th percentile price difference
as $P D_{90 \_10}=\frac{P^{90}-P^{10}}{\bar{P}}$, where $P^{90}$ and $P^{10}$ is the 90 th percentile and $10^{\text {th }}$ percentile of price, respectively. Table 4 summarizes the posted price dispersion in the sample.

Table 4: Summary statistics for posted price dispersion

|  | Mean | S.D | Min | Max |
| :--- | :--- | :--- | :--- | :--- |
| PD | 0.30 | 0.25 | 0 | 2.00 |
| CV | 0.13 | 0.11 | 0 | 1.36 |
| $\mathrm{PD}_{90 \_10}$ | 0.27 | 0.22 | 0 | 1.96 |
| N | 9,911 |  |  |  |

Note: 1. Data is extracted from Tmall (Site 1).
2. Each observation is a book. Price dispersion is calculated using posted prices, excluding shipping cost.
3. Books that are only available at one store are dropped.

A $30 \%$ percentage price difference and a $13 \%$ coefficient of variation are found. These numbers are very close to what is commonly suggested in the literature. The summary statistics on posted price discounts suggest that it is costly for consumers to conduct price comparison across stores, therefore, posted price dispersion is expected to increase with the number of stores. Figure 1 provides some evidence to support this prediction. The percentage price difference exhibits an increasing trend with the number of stores. The coefficient of variation and the percentage difference between the 90th and 10th prices are relatively more stable.


Figure 1: Average posted price dispersion by the number of stores

## 4 Hypotheses and Empirical Models

To formally test the relationship between price discount and the number of stores as well as the relationship between price dispersion and the number of stores, we propose two hypotheses and develop two econometric models to test our hypotheses.

### 4.1 Hypotheses

Standard economic theory believes that a greater degree of competition will lead to lower equilibrium prices. It is generally accepted that as the number of sellers increases, competition will become more intense. Accordingly, sellers have to reduce price to compete with rivals. This result is derived under the assumption of perfect information. If there is frictionless commerce in the internet market, increasing the number of firms would increase the effective degree of market competitiveness. This is confirmed by Clay et. al. [16] on the internet market where lower prices are found for books carried by more sellers. However, as it is demonstrated in our descriptive and summary analyses, sellers may conduct various strategies in information revelation and price promotions that would increase the cost for consumers to acquire product information and compare prices. Such strategies may reduce the effective degree of competition. To better understand the relationship between the number of sellers and competition in the internet market, we propose the following hypothesis. This hypothesis is developed based on the conventional economic result that the number of
firms is positively related to the degree of market competition and the assumption that the cost to obtain information on the internet is close to zero for the consumers.

## H1: Price discount decreases as the number of stores increases.

Previous literatures have documented the existence of online price dispersion [11], [16], [36]. They argued that although the development of Internet and other information technologies facilitate price search, it allows firms to adjust price frequently and adopt complicated price strategies. Consequently, consumer search cost may not reduce on the internet. When the number of sellers increases, the cost for consumers to review all sellers would increase. In addition, in online shopping, consumers face greater uncertainty in product and service quality. This allows larger sellers to take advantage of their established reputation to increase prices. Smaller sellers, on the other hand, tend to rely on purchases from inexperienced consumers and may offer deeper price cuts [29]. As the number of sellers increases, greater variation in the size and reputation among the sellers is likely to arise, which would lead to more extreme price strategies. We thus test the following hypothesis:

## H2: Price dispersion increases as the number of stores increases.

### 4.2 Empirical Models

We conduct two analyses to explore posted prices in response to competition and information asymmetry. We first study the relationship between price discount and the number of sellers. We then examine how posted price dispersion changes with the number of sellers.

### 4.2.1 Model 1: Price Discount Study

Let i denote book and j denote store. The following model is specified to test H 1 :

$$
\begin{equation*}
\text { Discount }_{\mathrm{ij}}=\alpha_{0}+\alpha_{1} \log \left(\text { Price }_{\mathrm{i}}\right)+\alpha_{2} \text { Sales }_{\mathrm{ij}}+\alpha_{3} \text { Ranking }_{\mathrm{j}}+\alpha_{4} \text { Market }_{\mathrm{i}}+\alpha_{5} \text { Listing }_{\mathrm{ij}}+\alpha_{6} \text { Store }_{\mathrm{j}}+\alpha_{7} \text { Review }_{\mathrm{ij}}+\varepsilon_{\mathrm{ij}} . \tag{1}
\end{equation*}
$$

In equation (1), the dependent variable Discount ${ }_{\mathrm{ij}}$ is the posted price discount of book i in store j . The explanatory variable market ${ }_{i}$ is the number of stores in the market for book $i$. Instead of using a continuous variable for the number of firms, we construct seven dummy variables: 2 stores, 3 stores, 4 stores, 5 stores, 6 - 10 stores, 11-20 stores and more than 20 stores. We incorporate them separately into the regressions to avoid any potential multicollinearity problem and more importantly, to capture any nonlinear impact of the number of sellers on price discounts. Price is the publisher's recommended price for book i . Sale $\mathrm{e}_{\mathrm{ij}}$ is the sales of book i in store j over the past month. Ranking $\mathrm{j}_{\mathrm{j}}$ is bookstores' sales ranks based on past month sales. Notice that Sale $\mathrm{ij}_{\mathrm{ij}}$ is measured in book-bookstore level while Ranking ${ }_{\mathrm{j}}$ is store specific. Listing $_{\mathrm{ij}}$ is a set of information about the book provided by the bookstores which includes the number of photos, whether the listing includes table of contents, editor's review, author information, and content summary. Store $_{j}$ is a list of store-level fees and promotions that consist of the shipping fee, free shipping, gift with purchase, voucher for future consumption, direct price deduction, and consumer credit reward. Some of these storelevel discounts are conditional on the total value of the purchase. Review $\mathrm{w}_{\mathrm{ij}}$ covers a set of measures for aggregate consumer review. Consumers may leave a review on the book after they make a purchase. We include the percentage of positive reviews by taking the ratio between total number of positive keywords and total number of reviews. We also include the percentage of negative reviews which is defined in a similar fashion. Additionally, Tmall reports consumer reviews on book organization and book design. Consumers may rate the organization of the book as comfortable, average, and difficult; Consumers may rate the design of the book as exquisite, pleasingly simple, and average. We include the percentage of reviews on organization and on design in the two preferred categories (average and comfortable for organization, and pleasingly simple and exquisite for design) in our model. In addition, book types and store provinces are also controlled for to capture unobserved demand and supply heterogeneities. Table 5 provides a detailed description of the explanatory variables included in model 1.

Table 5: Description of Explanatory Variables in Model 1

| Variable | Description |
| :--- | :--- |
| Log(Price) | Log of book $i$ 's publisher's recommended price (in Chinese Yuan) |
| Past Month Sales | Total sales of book $i$ in store $j$ in the past month |
| Seller Ranking (Past Month Sales) | Ranking of total sales of all books in the past month for store $j$ |
| 2 Stores | Dummy variable $=1$ if there are two stores offering book $i$ |
| 3 Stores | Dummy variable $=1$ if there are three stores offering book $i$ |
| 4 Stores | Dummy variable $=1$ if there are four stores offering book $i$ |
| 5 Stores | Dummy variable $=1$ if there are five stores offering book $i$ |
| $6-10$ Stores | Dummy variable $=1$ if there are between six and ten stores offering book $i$ |
| $11-20$ Stores | Dummy variable $=1$ if there are between eleven and twenty stores offering book $i$ |
| More than 20 Stores | Dummy variable $=1$ if there are more than twenty stores offering book $i$ |
| Number of Photos | Number of photos provided for book $i$ by store $j$ |

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718-1876 Electronic Version
VOL 16 / ISSUE 2 / MAY 2021 / 140-154
© 2021 Universidad de Talca - Chile

This paper is available online at www.jtaer.com
DOI: 10.4067/S0718-18762021000200110

| Table 5: continuation |  |
| :---: | :---: |
| Table of Contents (Catalog) | Dummy variable $=1$ if store $j$ provides table of contents for book $i$ |
| Editor's Review | Dummy variable $=1$ if store $j$ provides editor's review for book $i$ |
| Author Introduction | Dummy variable $=1$ if store $j$ provides author introduction for book $i$ |
| Content Summary | Dummy variable $=1$ if store $j$ provides content summary for book $i$ |
| Shipping Fee | Express shipping fee charged by store $j$ (in Chinese Yuan) |
| Free Shipping (< 50 RMB ) | Dummy variable $=1$ if store $j$ offers conditional free shipping based on the total amount of purchase in the order and the required amount is less than 50 Chinese Yuan |
| Free Shipping (50-100 RMB) | Dummy variable $=1$ if store $j$ offers conditional free shipping based on the total amount of purchase and the required amount is between 50 and 100 Chinese Yuan |
| Free Shipping (> 100 RMB) | Dummy variable $=1$ if store $j$ offers conditional free shipping based on the total amount of purchase and the required amount is more than 100 Chinese Yuan |
| Gift with Purchase ( $\leq 100 \mathrm{RMB}$ ) | Dummy variable $=1$ if store $j$ offers gifts based on the total amount of purchase and the required amount is less than 100 Chinese Yuan |
| Gift with Purchase (> 100 RMB ) | Dummy variable $=1$ if store $j$ offers gifts based on the total amount of purchase and the required amount is more than 100 Chinese Yuan |
| Voucher | Dummy variable $=1$ if store $j$ offers any amount of voucher for future purchases |
| Price Deduction ( $\leq 100 \mathrm{RMB}$ ) | Dummy variable $=1$ if store $j$ offers direct price reduction based on the total amount of purchase and the required amount is less than or equal to 100 Chinese Yuan |
| Price Deduction (> 100 RMB ) | Dummy variable $=1$ if store $j$ offers direct price reduction based on the total amount of purchase and the required amount is more than 100 Chinese Yuan |
| Credit Reward | Dummy variable $=1$ if store $j$ offers any credit reward |
| Percentage Positive Comments | Count of positive key words in the consumer reviews divided by the number of consumer review for book $i$ sold in store $j$ |
| Percentage Negative Comments | Count of negative key words in the consumer reviews divided by the number of consumer review for book $i$ sold in store $j$ |
| Organization-Average (\%) | Percentage of consumer reviews that rate the organization of the book as average for book $i$ sold in store $j$ |
| Organization-Comfortable (\%) | Percentage of consumer reviews that rate the organization of the book as comfortable for book $i$ sold in store $j$ |
| Design-Pleasantly Simple(\%) | Percentage of consumer reviews that rate the design of the book as pleasantly simple for book $i$ sold in store $j$ |
| Design-Exquisite(\%) | Percentage of consumer reviews that rate the design of the book as exquisite for book $i$ sold in store $j$ |

### 4.2.2 Model 2: Price Dispersion Study

In model 2, we study the relationship between posted price dispersion and the number of sellers (H2). Let i denote book. The model is specified as follows.

$$
\begin{equation*}
\text { Dispersion }_{i}=\beta_{0}+\beta_{1} \text { Log }\left(\text { Price }_{i}\right)+\beta_{2} \text { Market }_{i}+\beta_{3} B_{i}+\beta_{4} \text { Store_Variation } i+\varepsilon_{i} . \tag{2}
\end{equation*}
$$

The dependent variable Dispersion is the post price dispersion of book i. Similar to model (1), Market is a set of dummies that represent the number of stores and Price $i$ is the publisher's recommended price for book $\mathrm{i} . \mathrm{B}_{\mathrm{i}}$ includes book i's sales rank from the past month to capture each book demand. $\mathrm{B}_{\mathrm{i}}$ also includes a set of availability measures: whether book i is available in a top 10 selling store, a top 11-50 selling store, or a direct sales store. Additionally, dummy variables indicating whether book $i$ is available in stores with different promotions are also included in the model. Store_Variation ${ }_{i}$ consists of a set of measures on seller heterogeneity, which includes the standard deviation of the sales ranks of stores in the past month and the standard deviation of store-level consumer ratings on the match between seller description and the actual product, the overall service quality, and the shipping. Table 6 provides detailed descriptions of the explanatory variables included in model 2.

Table 6: Description of Explanatory Variables in Model 2

| Variable | Description |
| :--- | :--- |
| Log(Price) | Log of book $i$ 's publisher's recommended price (in Chinese Yuan) |
| 2 Stores | Dummy variable $=1$ if there are two stores offering book $i$ |
| 3 Stores | Dummy variable $=1$ if there are three stores offering book $i$ |
| 4 Stores | Dummy variable $=1$ if there are four stores offering book $i$ |
| 5 Stores | Dummy variable $=1$ if there are five stores offering book $i$ |
| $6-10$ Stores | Dummy variable $=1$ if there are between six and ten stores offering book $i$ |
| $11-20$ Stores | Dummy variable $=1$ if there are between eleven and twenty stores offering <br> book $i$ |
| Book Sales Ranking | Sales rank of book $i$ based on the total sales of the book across all stores <br> in the past month |
| Available at a Top 10 Bookstore | Dummy variable $=1$ if book $i$ is available in one or more of the top 10 <br> bookstores based on the past month's sales rank of the stores |

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718-1876 Electronic Version
VOL 16 / ISSUE 2 / MAY 2021 / 140-154
© 2021 Universidad de Talca - Chile

This paper is available online at www.jtaer.com
DOI: 10.4067/S0718-18762021000200110
Table 6: continuation

| Available at a Top 11-50 Bookstore | Dummy variable $=1$ if book $i$ is available in one or more of the to bookstores based on the past month's sales rank of the stores |
| :---: | :---: |
| Available at a Direct Bookstores | Dummy variable $=1$ if book $i$ is available in its own publisher's direct sales online store |
| Available at a Bookstore w | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer conditional free shipping based on the total amount of purchase and the required amount is less than 50 Chinese Yuan |
|  | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer conditional free shipping based on the total amount of purchase and the required amount is between 50 and 100 Chinese Yuan |
| valable at a Bookstore w e shipping (> 100 RMB) | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer conditional free shipping based on the total amount of purchase and the required amount is greater than 100 Chinese Yuan |
|  | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer gifts based on the total amount of purchase and the required amount is less than 100 Chinese Yuan |
|  | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer gifts based on the total amount of purchase and the required amount is greater or equal to 100 Chinese Yuan |
| vailable at a Bookstore ucher | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer any voucher for future purchases |
| Price Deduction (< 100 RMB) |  |
| Price Deduction ( $\geq 100$ RMB) | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer direct price reduction based on the total amount of purchase and the required amount is greater than or equal to 100 Chinese Yuan |
| Credit Award | Dummy variable $=1$ if book $i$ is available in one or more of stores that offer credit reward |
| Variation of Bookstore (Past Month) | Standard deviation of sales rankings of stores offering book $i$ based on bookstore level total sales |
| Variation of Bookstore Ratings (Description Match) | Standard deviation of store-level consumer review scores on the match between the actual product and the description provided by the stores |
| Variation of Bookstore Ratings (Service) | Standard deviation of store-level consumer review scores on the overall service quality provided by the stores |
| ariation of Bookstore Ratings Shipping) | Standard deviation of store-level consum service provided by the stores |

## 5 Results

In this section, we present the regression results of our empirical analyses based on equation (1) and equation (2). We provide an interpretation of our estimates and discuss the implications from our results.

### 5.1 Result 1: Price Discount Study

Table 7 presents the regression results of model 1 for the full listing sample, the sample of books with positive sales in the past month, and the sample of books with zero sales in the past month.

Table 7: Estimated results for posted price discounts

|  | Full Sample | Books with Positive Sales | Books with Zero Sales |
| :--- | :--- | :--- | :--- |
| Log(Price) | $0.06^{* * *}$ | $0.08^{* * *}$ | $0.02^{* * *}$ |
|  | $(24.60)$ | $(30.22)$ | $(4.15)$ |
| Past Month Sales $\left(\times 10^{-6}\right)$ | 14.20 | $2.44^{* * *}$ |  |
|  | $(1.54)$ | $(2.65)$ | $-0.35^{* * *}$ |
| Seller Ranking <br> Month Sales $)\left(\times 10^{-3}\right)$ | $-0.16^{* * *}$ | $(-15.67)$ | $(-3.53)$ |
|  | Market Competitiveness |  |  | $-0.51)$ |
| 2 Stores $\left(\times 10^{-3}\right)$ | $6.08^{*}$ | $22.31^{* * *}$ | -1.83 |
|  | $(2.13)$ | $(3.21)$ | $(-0.68)$ |
| 3 Stores $\left(\times 10^{-3}\right)$ | 3.69 | $17.76^{* * *}$ | -3.87 |
|  | $(1.29)$ | $(2.74)$ | $(-1.37)$ |

Journal of Theoretical and Applied Electronic Commerce Research

This paper is available online at www.jtaer.com
DOI: 10.4067/S0718-18762021000200110

Table 7: continuation

| Table 7: continuation |  |  |  |
| :---: | :---: | :---: | :---: |
| 4 Stores ( $\times 10^{-3}$ ) | 1.24 | 14.07** | -5.17* |
|  | (0.43) | (2.21) | (-1.75) |
| 5 Stores ( $\times 10^{-3}$ ) | -1.94 | 11.04* | -8.95*** |
|  | (-0.65) | (1.72) | (-2.82) |
| 6-10 Stores ( $\times 10^{-3}$ ) | -9.41*** | 1.68 | -18.75*** |
|  | (-3.49) | (0.28) | (-6.49) |
| 11-20 Stores ( $\times 10^{-3}$ ) | -6.91*** | 0.37 | -16.55** |
|  | (-2.46) | (0.06) | (-5.41) |
| $\begin{aligned} & \text { More than } 20 \text { Stores } \\ & \left(\times 10^{-3}\right) \end{aligned}$ | -11.43*** | -0.85 | $-27.44^{* * *}$ |
|  | (-3.71) | (-0.14) | (-7.98) |
| Listing Information |  |  |  |
| Number of Photos | 0.02*** | 0.01*** | 0.08*** |
|  | (6.65) | (3.68) | (6.98) |
| Table of Contents(Catalog) | 0.04*** | 0.04*** | 0.03*** |
|  | (20.23) | (14.95) | (10.40) |
| Editor's Review | -0.01*** | -0.01*** | -0.02*** |
|  | (-5.48) | (-3.66) | (-5.47) |
| Author Introduction | -0.01*** | -0.01*** | -0.003 |
|  | (-6.79) | (-6.69) | (-1.45) |
| Content Summary | 0.01*** | 0.01*** | 0.02*** |
|  | (11.74) | (7.36) | (10.66) |

## Store-Level Fee and Discounts

| Shipping Fee ( $\times 10^{-3}$ ) | 3.56** | 3.66 | 3.64* |
| :---: | :---: | :---: | :---: |
|  | (2.09) | (1.30) | (1.65) |
| $\begin{aligned} & \text { Free Shipping (< } 50 \text { RMB) } \\ & \left(\times 10^{-3}\right) \end{aligned}$ | -2.53** | -13.76*** | 8.13*** |
|  | (-2.14) | (-7.21) | (4.95) |
| Free Shipping (50-100$\text { RMB) }\left(\times 10^{-3}\right)$ | 9.15*** | 4.31 | 19.35*** |
|  | (3.53) | (0.95) | (6.34) |
| $\begin{aligned} & \text { Free Shipping (> } 100 \\ & \text { RMB) } \end{aligned}$ | 0.01 | -0.04 | 0.16*** |
|  | (0.44) | (-1.29) | (7.71) |
| Gift with Purchase ( $\leq 100$ RMB) | 0.13*** | 0.11*** | 0.14*** |
|  | (14.83) | (6.84) | (12.47) |
| Gift with Purchase (> 100 | -0.02*** | -0.04*** | -0.01 |
|  | (-2.96) | (-3.12) | (-1.18) |
| Voucher | 0.03*** | 0.03*** | 0.02*** |
|  | (12.19) | (7.14) | (5.65) |
| $\begin{aligned} & \text { Price Deduction ( } \leq 100 \\ & \text { RMB) } \end{aligned}$ | 0.08*** | 0.06*** | $0.12^{* * *}$ |
|  | (4.57) | (2.85) | (4.48) |
| $\begin{aligned} & \text { Price Deduction (> } 100 \\ & \text { RMB) } \end{aligned}$ | -0.05*** | -0.02 | -0.10*** |
|  | (-3.16) | (-1.39) | (-3.36) |
| Credit Reward | -0.08*** | -0.09*** | -0.06*** |
|  | (-9.06) | (-6.20) | (-5.42) |

## Consumer Reviews

| Percentage <br> Comments | Positive | $0.43^{* * *}$ | $0.39^{* * *}$ |
| :--- | :--- | :--- | :--- |
|  | $(5.43)$ | $(5.25)$ |  |
| Percentage <br> Comments | $-1.51^{* * *}$ | $-1.29^{* * *}$ |  |
|  | $(-4.13)$ | $(-3.79)$ |  |
| Organization-Comfortable <br> $(\%)\left(\times 10^{-3}\right)$ | $(-0.63)$ | -2.70 |  |
|  | -0.23 | $(-0.01)$ |  |
| Design-Exquisite(\%) | $0.01^{* * *}$ | 0.01 |  |
|  | $(2.62)$ | $(0.33)$ |  |
| Book Type | -0.01 | $0.01^{* * *}$ |  |
| Store Province | $(1.35)$ | $(2.70)$ | Yes |
| N | Yes | -0.01 | Yes |
| Number of Book Titles | Yes | $(1.45)$ | 35,129 |
| $\mathrm{R}^{2}$ | 68,003 | Yes | 9,600 |

Note: 1. Data is collected from Tmall (Site 1).
2. Robust standard errors are reported in parentheses. ${ }^{*} p<0.1 ;{ }^{* *} p<0.05 ;{ }^{* * *} p<0.01$.

The results suggest a complicated relationship between the number of sellers and price discount. For books with positive sales, there is a substantial increase in price discount when the market changes from monopoly to duopoly; however, as more firms enter the market, the posted price discounts actually decrease, in other words, prices increase. H 1 is rejected. More strikingly, when the number of firms exceeds five, no significant difference is found between prices in larger market and prices in the monopoly market. For books without sales, the coefficients on 2 stores and 3 stores are not statistically significant. But as the number of stores increases, the magnitude of the coefficients rises and the significance level also improves. To summarize, for popular books, the initial introduction of competition causes prices to drop, but more entrants indeed weaken competition. This is because as more sellers enter the market, the cost for consumers to conduct search increases, which deters consumer search. Consumers therefore are not expected to search over all the stores and their purchasing decision become more random. The return for sellers to cut price indeed decreases as more rivals enter the market. Hence, instead of lowering prices, sellers increase prices when facing more competitors.

Posted price discounts increase with the value of the book. A one percentage increase in the publisher's recommended price leads to an $8 \%$ increase in discounts for books with positive sales and a $2 \%$ increase in discounts for books with zero sales. Shipping charges, though not directly absorbed by the sellers, influence sellers' pricing decisions. Our result suggests that there is a statistically significant impact of shipping costs only on books without any sales. The estimated effect, however, is of very small magnitude.

Store-level discounts can serve three purposes. Sellers may provide illusionary discounts by offering store-wide promotions while increasing individual prices; new sellers may use store-level promotions to attract curious consumers and gradually build up loyal consumer base; sellers may also want to clear inventory by adopting a semi loss-leader pricing where less popular items are given double discounts while more popular products experience price increases. As it is demonstrated in Table 5, when free shipping is offered for a small purchase, sellers decrease the discount for popular books while increase the discount for less popular books. Free shipping based on larger purchases does not affect popular books significantly but continue to lead to larger discounts for less popular books. These support the inventory clearing theory. Direct price deductions are found to have a positive effect on percentage price discount for small purchases but a negative effect for larger purchases. Notice that the coefficients in the zero-sale sample are much larger than the coefficients in the positive sales sample. Voucher and gift with purchase promotions based on a small purchases are associated with higher discounts. This is because vouchers are given out with the expectation of future consumptions, and gifts usually serve as a courtesy and an attempt to build a good impression of the sellers.

Unlike information such as pictures and author introduction that sellers voluntarily reveal, consumer reviews are passive private information disclosure. Consumer reviews usually provide information about product and service qualities based on their actual experience, and therefore, are perceived to have greater credibility by future consumers. Firms benefit from positive consumer reviews and are often wary of poor reviews. On Tmall, each listing page puts the total number of reviews on the top of the page which suggests its importance. This number of review also serves as an indication of the popularity of the store and the book. Additionally, consumer reviews are good indications of the future demand of the book. Seller can adjust prices of the books based on the reviews. Intuitively, one would expect that positive comments increase prices while negative comments drive down prices. Our results however, suggest the opposite. A one percentage increase in positive comments leads to a $0.39 \%$ increase in percentage price discounts; a one percentage increase in negative comments leads to a $1.29 \%$ decrease in percentage price discounts for books with positive sales. When books receive good reviews, which are likely to lead to higher demand, rather than increasing prices, sellers offer even greater discounts on these books. When books receive negative comments which are likely to decrease their demand, instead of lowering prices, sellers increase the prices of these books. One explanation is provided by Berger et al. [7] that negative publicity may still increase sales. Another explanation is that sellers conduct loss-leader pricing strategy. They lower prices of popular products to attract traffic into their stores hoping that they would also purchase other less popular products with higher mark-up.

The study on posted price discounts adds novel findings to the literature. Prices do not always decrease with the number of firms even on the internet market where price information is easy to obtain. For popular books, competition only strengthens when a market moves from monopoly to a duopoly. When the market further expands, competition effect is mitigated and prices eventually move back to monopoly level.

### 5.2 Result 2: Price Dispersion Study

Regression analyses for equation 2 are conducted for three measures of price dispersion, percentage price difference, coefficient of variation, and the percentage price difference between the 90th and 10th percentile prices. Books that are only carried by one store are not included in the price dispersion study.

As can be seen in Table 8, the value of the book is not a significant determinant of posted price dispersion. This is consistent with Zhao et. al. [48] which suggest that sellers' inability of incorporating consumer search behavior leads to the irrelevance between price level and posted price dispersion.

Posted price dispersion increases as the number of stores increases. H2 is confirmed. Compared to books that are carried by more than 20 stores, books with only two sellers have a $44 \%$ less percentage price difference and $5 \%$ less coefficient of variation. The change in the percentage price difference is substantially higher than the change in the

Journal of Theoretical and Applied Electronic Commerce Research

This paper is available online at www.jtaer.com
DOI: 10.4067/S0718-18762021000200110
coefficient of variation. As the number of stores increases, some stores charge extreme prices while the other stores have relatively stable prices. The positive relationship between the number of firms and posted price dispersion arises largely from the existence of extreme prices. The level of variation, however, stabilizes much more quickly.

Table 8: Estimated results for posted price dispersion

|  | PD | CV | $\mathrm{PD}_{90 \_10}$ |
| :---: | :---: | :---: | :---: |
| Log(Price) | 0.01 | 0.01 | 0.01 |
|  | (1.17) | (1.48) | (0.79) |
| 2 Stores | -0.44*** | -0.05*** | -0.20*** |
|  | (-22.75) | (-7.37) | (-13.21) |
| 3 Stores | -0.35*** | -0.03*** | -0.11*** |
|  | (-18.04) | (-5.02) | (-7.30) |
| 4 Stores | -0.30*** | -0.03*** | -0.06*** |
|  | (-15.47) | (-4.24) | (-3.94) |
| 5 Stores | -0.25*** | -0.02*** | -0.01 |
|  | (-12.91) | (-2.83) | (-0.54) |
| 6-10 Stores | -0.19*** | -0.01*** | 0.04*** |
|  | (-10.56) | (-2.47) | (3.21) |
| 11-20 Stores | -0.08*** | 0.00 | 0.01 |
|  | (-4.55) | (0.39) | (1.08) |
| Book Sales Ranking ( $\times 10^{-3}$ ) | -0.59*** | -0.18*** | -0.43** |
|  | (-5.53) | (-4.42) | (-4.99) |
| Available at a Top 10 Bookstore | -0.07*** | -0.03*** | -0.06*** |
|  | (-8.48) | (-9.34) | (-9.09) |
| Available at a Top 11-50 Bookstore | -0.01 | -0.01** | -0.02** |
|  | (-1.44) | (-2.18) | (-2.31) |
| Available at a Direct-Sale Bookstores | 0.02*** | 0.01*** | 0.02*** |
|  | (2.56) | (2.89) | (3.22) |
| Available at a Bookstore with Free shipping (< 50 RMB ) | -0.04*** | -0.03*** | -0.04*** |
|  | (-6.86) | (-8.49) | (-7.62) |
| Available at a Bookstore with Free shipping (50-100 RMB) | -0.04*** | -0.02*** | -0.04*** |
|  | (-6.29) | (-7.74) | (-6.63) |
| Available at a Bookstore with Free shipping (> 100 RMB) | -0.11** | -0.05*** | -0.08* |
|  | (-1.99) | (-2.40) | (-1.87) |
| Available at a Bookstore with Gift with Purchase ( < 100 RMB) | 0.17*** | 0.07*** | 0.14*** |
|  | (6.55) | (7.63) | (7.48) |
| Available at a Bookstore with Gift with Purchase ( $\geq 100$ RMB) | -0.06*** | -0.03*** | -0.09*** |
|  | (-2.88) | (-4.55) | (-5.30) |
| Available at a Bookstore with Voucher | -0.06*** | -0.03*** | -0.06*** |
|  | (-9.33) | (-10.72) | (-10.49) |
| Available at a Bookstore with Price Deduction (< 100 RMB) | 0.02 | 0.02 | 0.03 |
|  | (0.31) | (1.23) | (0.82) |
| Available at a Bookstore with Price Deduction ( $\geq 100$ RMB) | 0.08 | 0.01 | 0.03 |
|  | (1.30) | (0.68) | (0.67) |
| Available at a Bookstore with Credit Award | 0.09*** | 0.04*** | 0.09*** |
|  | (6.42) | (7.10) | (8.34) |
| Variation of Bookstore Rankings (Past Month) ( $\times 10^{-3}$ ) | 0.30*** | 0.19*** | 0.30*** |
|  | (3.63) | (3.97) | (3.80) |
| Variation of Bookstore Ratings (Description Match) | $0.41^{* * *}$ | 0.24*** | $0.38{ }^{* * *}$ |
|  | (3.96) | (4.22) | (3.80) |
| Variation of Bookstore Ratings(Service) | -0.87*** | -0.48*** | -0.85*** |
|  | (-8.84) | (-8.25) | (-8.80) |
| Variation of Bookstore Ratings(Shipping) | 0.20*** | 0.10** | 0.21** |
|  | (2.70) | (2.18) | (2.82) |
| Book Type | Yes | Yes | Yes |
| N | 9,911 | 9,911 | 9,911 |
| $\mathrm{R}^{2}$ | 0.26 | 0.08 | 0.18 |

Note: 1. Data is collected from Tmall (Site 1).
2. $z$-statistics are reported in parentheses. ${ }^{*} p<0.10 ;{ }^{* *} p<0.05 ;{ }^{* * *} p<0.01$.

Aggregate sales of stores indicate its reputation in the market and existing consumer base. Stores of different sizes thus have different market power. Hence, the availability at different bookstores could lead to different degree of price dispersion for a book. The estimated coefficients on the availability at a top 10 and a top 11-50 bookstore are both
negative. The coefficients on the availability of at a top 10 store are significant at the 0.01 level while the coefficients on the availability at a top 11-50 store are marginally significant. The existence of a large competitor increases market competition and drives down price dispersion.

The search cost for store-level promotions is substantially lower than the search cost for item-specific discounts. Thus, store-level discounts usually draw substantially more traffic to the stores compared to individual price cuts. When store promotions are observed, rival stores may also conduct store-wide discounts or lower individual item prices, especially when they share a large selection of identical products. Store-level promotions, therefore, foster competition and drive price convergence. This is supported by our estimated results, except for gifts based on small purchases and credit rewards. But as discussed in Section 4.1, gifts with purchase usually serve as to build good impression and connection with consumers rather than price reduction.

The effect of the heterogeneity in seller sales rankings is positive and statistically significant. The variations in consumer ratings, however, have different effects. Heterogeneity in consumers' perception on the match between sellers' description and the actual products as well as the shipping speed foster posted price dispersion while heterogeneity in service attitude ratings leads to lower dispersion.

The analyses on posted price dispersion find consistent results with the literature that dispersion increases with the number of firms. Stores face different levels of competition due to past sales. The existence of a large store fosters competition and leads to lower dispersion, but a market with larger store heterogeneity would allow stores to charge different prices.

## 6 Conclusion

Using fiction book listing prices and other book level and store level information from Tmall (Site 1), the largest online Business-to-Consumer (B2C) platform in China, this paper examines the complicated information structure on the internet market and its impact on the competitiveness of the market. Contrary to the prediction from standard economic theory that competition would increase as the number of firms increases in the market, leading to lower price and price dispersion, we find that on the internet market, the result is more complex. With sellers' active use of various pricing strategies and promotions, an increase in the number of firms lead to an increase in the cost for consumers to conduct search, and subsequently a deterrence of consumer search and a higher level of information asymmetry. Entry of firms only strengthens competition when there is a small number of incumbents on the market. Specifically, we find that competition effect only dominates and leads to price reduction when there are less than two sellers in the market. As the market expands, consumers have difficulty searching across stores and comparing and processing information. The expected benefits for sellers to lower price in order to attract consumers are reduced. Thus, having recognized consumers' inability to search all the stores, rather than reducing prices, sellers would raise prices. While we confirm that information asymmetry reduces the effective degree of competition [40], [42], [43], we find empirical evidence to demonstrate that increasing the number of firms would not always improve market competitiveness in the presence of information asymmetry [44]. Existing studies that show such empirical evidence focus on physical retail markets. For example, Perloff et al. [38] find higher prices due to new entries in retail drug stores and Thomadsen [45] find such pricing pattern in the fast food industry. To the best of our knowledge, we are the first to find a non-monotonic relationship between price and the number of firms in the online market.

While internet facilitates consumers to obtain product and price information at low cost, internet also allows sellers to conduct numerous forms of information revelation and promotional strategies. The amount of information may serve to obfuscate consumers, rather than help consumers in decision making. In addition to the flexibility of price adjustment [9], we show that the different level of information revealed by sellers as well as the many promotions offered by sellers further drive incomplete consumer search. While competing for consumers, sellers have indeed collectively mitigated the degree of competition. This allows them to charge higher prices even when facing large number of rivals. Consumers on the other hand, though have access to larger amount of information at lower cost and can receive many forms of discounts from sellers, may find it too costly to process the information.

To ensure that consumer benefit from the availability of information on the internet market, certain forms of standardization could be employed by online marketplace. Instead of allowing retailers to design their own page for each product and provide relevant product descriptions, online platforms may create a standardized page for each product with all the product information. The available sellers and their current prices will also be listed in each product page. Other brief seller information, such as aggregate consumer rating and aggregate sales may also be presented, depending on the availability of space. If consumers decide that the product is of interest, they will only need to compare the sellers before they make a purchase. Indeed, Amazon (Site 2 ) uses similar web design in their book market. Such web design reduces consumer search cost from the different product descriptions presented by retailers, which further limits sellers' flexibility to change price. In the meantime, online marketplace may also create additional mechanism to detect any high frequency price changes that are deliberately used by retailers to confuse consumers. If such behaviors are found, retailers would receive warning or punishment depending on the extent of their behaviors. These controls and measures would be useful to allow consumers to benefit from the vast information on the internet market, rather than be harmed by the availability of information.

Our empirical analyses rely on cross-sectional data which assume that prices are in equilibrium stage. Sellers may adjust prices as competitors enter or exit the market, thus our data may not be able to address the dynamics of online prices. For future study, data for multiple time periods may be collected and the present study may be extended using panel data regression techniques. It would be interesting to explore the effect of new entrants on the incumbent sellers' prices as well as the exit of competitors on remaining sellers' pricing decisions.

## Acknowledgments

We would like to thank Xinwei Jiang and Kuangyu Chen for their generous assistance in the data collection process. We also want to thank Kenny Cheng, Steven Slutsky, Jonathan Hamtilon, and Chunrong Ai for their valuable comments and advice. We appreciate Roger Blair, Thomas Knight, Scott Kostyshak, Richard Romano, Mark Rush, and David Sappington for their insightful suggestions in the economics department seminar. We are also very grateful to the seminar participants at Loyola University Chicago, Econ One Research, St. Lawrence University, and the European Association for Research in Industrial Economics for their excellent questions and suggestions. Financial support from the National Natural Science Foundation of China (71702109), the Research Foundation of Department of Education of Guangdong Province (2018WTSCX124), and the Philosophy and Social Science Foundation of Shenzhen (SZ2020C009) are gratefully acknowledged.

## Website List

## Site 1: Tmall <br> http://www.tmall.com

Site 2: Amazon
http://www.amazon.com

## References

[1] J. Alba, J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, and S. Wood, Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces, Journal of Marketing, vol. 61, no. 3, pp. 38-53, 1997.
[2] F. Ancarani and V. Shankar, Price levels and price dispersion within and across multiple retailer types: Further evidence and extension, Journal of the Academy of Marketing Science, vol. 32, no. 2, pp. 176-187, 2004.
[3] S. Ba, J. Stallaert and Z. Zhang, Research note-Online price dispersion: A game-theoretic perspective and empirical evidence, Information Systems Research, vol. 23, no. 2, pp. 575-592, 2012.
[4] J.P. Bailey, Electronic commerce: Prices and consumer issues for three products: Books, compact discs and software, OECD Digital Economy Papers 32, OECD Publishing, Paris, 1998.
[5] J.Y. Bakos, Reducing buyer search costs: Implications for electronic marketplaces, Management Science, vol. 43, no. 12, pp. 1676-1692, 1997.
[6] M. R. Baye, J. Morgan and P. Scholten, Price dispersion in the small and in the large: Evidence from an internet price comparison site, Journal of Industrial Economics, vol. 52, no. 4, pp. 463-496, 2004.
[7] J. Berger, A. T. Sorensen and S. J. Rasmussen, Positive effects of negative publicity: When negative reviews increase sales, Marketing Science, vol. 29, no. 5, pp. 815-827, 2010.
[8] T. Blake, S. Moshary, K. Sweeney, and S. Tadelis, Price salience and product choice, National Bureau of Economic Research, Cambridge, MA, Working Paper 25186, 2018.
[9] J. Boivin, R. Clark and N. Vincent, Virtual borders, Journal of International Economics, vol. 86, no. 2, pp. 327335, 2012.
[10] J. R. Brown and A. Goolsbee, Does the internet make markets more competitive? Evidence from the life insurance industry, Journal of Political Economy, vol. 110, no. 3, pp. 481-507. 2002.
[11] E. Brynjolfsson and M. D. Smith, Frictionless commerce? A comparison of internet and conventional retailers, Management Science, vol. 46, no. 4, pp. 563-585, 2000.
[12] A. Cavallo, B. Neiman and R. Rigobon, Currency unions, product introductions, and the real exchange rate, Quarterly Journal of Economics, vol. 129, no. 2, pp. 529-595, 2014.
[13] A. Chandra and M. Tappata, Consumer search and dynamic price dispersion: An application to gasoline markets, The RAND Journal of Economics, vol. 42, no. 4, pp. 681-704, 2011.
[14] Y. Chen, G. Iyer and A. Pazgal, Limited memory, categorization, and competition, Marketing Science, vol. 29, no. 4, pp. 650-670, 2010.
[15] J. Chevalier and A. Goolsbee, Measuring prices and price competition online: Amazon.com and BarnesandNoble.com, Quantitative Marketing and Economics, vol. 1, no. 2, pp. 203-222, 2003.
[16] K. Clay, R. Krishnan and E. Wolff, Prices and price dispersion on the web: Evidence from the online book industry, Journal of Industrial Economics, vol. 49, no. 4, pp. 521-539, 2001.
[17] K. Clay, R. Krishnan, E. Wolff, and D. Fernandes, Retail strategies on the web: Price and non-price competition in the online book industry, Journal of Industrial Economics, vol. 50, no. 3, pp. 351-367, 2002.
[18] E. K. Clemons, I. H. Hann and L.M. Hitt, Price dispersion and differentiation in online travel: An empirical investigation, Management Science, vol. 48, no. 4, pp. 534-549, 2002.
[19] P. A. Diamond, A model of price adjustment, Journal of Economic Theory, vol. 3, pp.156-168, 1971.
[20] M. Dinerstein, L. Einav, J. Levin, and N. Sundaresan, Consumer price search and platform design in internet commerce, American Economic Review, vol. 108, no. 7, pp.1820-1859, 2018.
[21] G. Ellison and S. Ellison, Search, obfuscation, and price elasticities on the internet, Econometrica, vol. 77, no. 2, pp. 427-452, 2009.
[22] S. Erevelles, E. Rolland and S. Srinivasan, Are prices really lower on the internet? An analysis of the vitamin industry, Working Paper, University of California, Riverside, 2001.
[23] Y. Gorodnichenko, V. Sheremirov and O. Talavera, Price setting in online markets: Does it click? Journal of the European Economic Association, vol.16, no.6, pp. 1764-1811, 2018.
[24] E. Honka, Quantifying search and switching costs in the US auto insurance industry, The RAND Journal of Economics, vol. 45, no. 4, pp. 847-884, 2014.
[25] A. Hortaçsu and C. Syverson, Product differentiation, search costs, and competition in the mutual fund industry: A case study of S\&P 500 index funds, Quarterly Journal of Economics, vol. 119, no. 2, pp. 403-456, 2004.
[26] T. Hossain and J. Morgan, Plus shipping and handling: Revenue (non) equivalence in field experiments on eBay, Advances in Economic Analysis and Policy, vol. 5, no. 2, 2006.
[27] M. C. W. Janssen and José, Luis Moraga-González, Strategic pricing, consumer search and the number of firms, Review of Economic Studies, vol. 71, no. 4, pp. 1089-1118, 2004.
[28] L. Kutlu, Limited memory consumers and price dispersion, Review of Industry Organization, vol. 46, no. 4, pp. 349-357, 2015.
[29] G. Lewis, Asymmetric information, adverse selection and online disclosure: The case of eBay motors, American Economic Review, vol. 101, no. 4, pp. 1535-1546, 2011.
[30] J. Lindsey-Mullikin and D. Grewal, Imperfect information: The persistence of price dispersion on the web, Journal of the Academy of Marketing Science, vol. 34, no. 2, pp. 236-243, 2006.
[31] P. Lünnermann and L. Wintr, Price stickiness in the U.S. and Europe revisited: Evidence from internet prices, Oxford Bulletin of Economics and Statistics, vol. 73, no. 5, pp. 593-621, 2011.
[32] J. G. Lynch and D. Ariely, Wine online: Search costs affect competition on price, quality, and distribution, Marketing Science, vol. 19, no. 1, pp. 83-103, 2000.
[33] F.S. Morton, F. Zettelmeyer and J. Silva-Risso, Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities? Quantitative Marketing and Economics, vol. 1, no. 1, pp. 65-92, 2003.
[34] E. Orlov, How does the internet influence price dispersion? Evidence from the airline industry, Journal of Industrial Economics, vol. 59, no. 1, pp. 21-37, 2011.
[35] X. Pan, B. T. Ratchford and V. Shankar, Can price dispersion in online markets be explained by differences in etailer service quality? Journal of the Academy of Marketing Science, vol. 30, no. 4, pp. 433-445, 2002.
[36] X. Pan, B. T. Ratchford and V. Shankar, Price dispersion on the internet: A review and directions for future research, Journal of Interactive Marketing, vol. 18, no. 4, pp. 116-135, 2004.
[37] B. K. Pathak, Comparison shopping agents and online price dispersion: A search cost based explanation, Journal of Theoretical and Applied Electronic Commerce Research, vol. 7, no. 1, pp. 64-76, 2012.
[38] J. Perloff, V. Suslow and P. Seguin, Higher Prices from Entry: Pricing of Brand-Name Drugs. Mimeo/University of California, Berkeley, 2006.
[39] P. Pereira, Do lower search costs reduce prices and price dispersion? Information Economics and Policy, vol. 17, no. 1, pp. 61-72, 2005.
[40] S. C. Salop, Information and monopolistic competition, American Economic Review, vol. 66, pp. 240-245, 1976.
[41] B. D. L. Santos, A. Hortaçsu and M. R. Wildenbeest, Testing models of consumer search using data on web browsing and purchasing behavior, American Economic Review, vol. 102, no. 6, pp. 2955-2980, 2012.
[42] T. Scitovsky, Ignorance as a source of oligopoly power, American Economic Review, vol. 40, pp. 48-53, 1950.
[43] J.E. Stiglitz, Equilibrium in Product Markets with Imperfect Information, American Economic Review, vol. 69, no.2, pp. 339-345, 1979.
[44] J.E. Stiglitz, Competition and the number of firms in a market: Are duopolies more competitive than atomistic markets? Journal of Political Economy, vol. 96, no. 5, pp. 1041-1061, 1987.
[45] R. Thomadsen, Product positioning and competition: the role of location in the fast food industry, Marketing Science, vol. 26, no. 6, pp. 792-804, 2007.
[46] K. L. Wakefield and J. J. Inman, Who are the price vigilantes? An investigation of differentiating characteristics influencing price information processing, Journal of Retailing, vol. 69, no. 2, pp. 216-233, 1993.
[47] J. Waldfogel and L. Chen, Does information undermine brand? Information intermediary use and preference for branded web retailers, Journal of Industrial Economics, vol. 54, no. 4, pp. 425-449, 2006.
[48] K. Zhao, X. Zhao and J. Deng, Online price dispersion revisited: How do transaction prices differ from listing prices? Journal of Management Information Systems, vol. 32, no. 1, pp. 261-290, 2015.

